Embedding median graphs into minimal distributive $\vee$-semi-lattices

Alain Gély, Miguel Couceiro, Amedeo Napoli

Embedding median graphs into minimal distributive ∨-semi-lattices

Alain Gély[1], Miguel Couceiro[2], and Amedeo Napoli[2]

[1] Université de Lorraine, CNRS, LORIA, F-57000 Metz, France
[2] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
{alain.gely,miguel.couceiro,amedeo.napoli}@loria.fr

Abstract. It is known that a distributive lattice is a median graph, and that a distributive ∨-semi-lattice can be thought of as a median graph iff every triple of elements such that the infimum of each couple of its elements exists, has an infimum. Since a lattice without its bottom element is obviously a ∨-semi-lattice, using the FCA formalism, we investigate the following problem: Given a semi-lattice L obtained from a lattice by deletion of the bottom element, is there a minimum distributive ∨-semi-lattice L_d such that L can be order embedded into L_d? We give a negative answer to this question by providing a counter-example.

Keywords: Median graph · Distributive lattice · order embedding· Formal Concept Analysis.

1 Motivation

Lattices and median graphs are two structures with many applications, in particular in classification and knowledge discovery. Median graphs are especially used in biology, for example in phylogeny, for modeling inter-species filiations. In phylogeny, one of the main problems is to find evolution trees for representing existing species from accessible DNA fragments. When several trees are leading to the same inter-species filiations, the preferred ones are the most "parsimonious", where the number of modifications such as mutations for example, is minimal for the considered species. However, several possible parsimonious trees may exist simultaneously. Such a situation arises with inverse or parallel mutations, e.g., when a gene goes back to a previous state or the same mutation appears for two non-linked species. This calls for a generic representation of such a family of trees.

Bandelt *et al.* [2,3] propose the notion of *median graph* to overcome this issue, since it was noticed that a median graph may encode all parsimonious trees. It is known that median graphs are related to lattices (see, e.g., [1,2]). Any distributive lattice is a median graph, and any median graph can be thought of as a distributive ∨-semi-lattice such that for all x, y, z such that the supremum of ech pair exists, then the supremum $\{x, y, z\}$ also exists.

Formal Concept Analysis (FCA) is based on lattice theory and can be used in classification and knowledge discovery. Uta Priss [13,14] made a first attempt to use the algorithmic machinery of FCA and the links between distributive lattices and median graphs, to analyze phylogenetic trees. However, not every concept lattice is distributive, and thus FCA alone does not necessarily outputs median graphs. A transformation should be designed to build a median graph from a concept lattice. In [14] Uta Priss sketches an algorithm to convert any lattice into a median graph. The key step is to transform any lattice into a distributive lattice. However, how to transform a lattice into a distributive one is not detailled in these papers.

In [4], Bandelt uses a data set from [15] to illustrate and evaluate median graph. In this introduction, we will re-use it to show the differences between median graph and FCA approaches. The example is an extract of mitochondrial DNA for 15 Kung individuals from a Khoisan-speaking hunter-gathered population in southern Africa. For some sequences in mitochondrial DNA (nucleotide positions, denoted by a, b, ... , j in table), a binary information indicates if a group of individuals owns the consensus version of the sequence (blanck value) or a variation for this sequence ($\times$ value). Eight individual groups are studied because some individuals share the same variants. For example, group 0 stands for 4 similar individuals in [4]. Individual group with no variation on any nucleotide positions (consensus group) is not shown on the table. These data are shown in Fig. 1 (Upper Left).

For these data, the median graph is shown in Fig. 1 (upper right). Vertices are either individual group (numbered from 0 to 7) or latent vertices, added such that from a group to an adjacent one, only one variation exists for nucleotide sequences (parcimony principle supposes that there is no chance that two variations arise exactly at the same moment for the same population in evolutionary process). This variation is indicated on edges. As an example, from consensus group to $L4$, the only variation occurs in sequence k, from $L4$ to 0 the only variation occurs in sequence j. As stated, this graph contains every parcimonious tree as covering tree. Median graph owns others good properties: remove edges labeled with a sequence variation produces two disconnected parts. One correspond to individuals with the variation, the other without the variation.

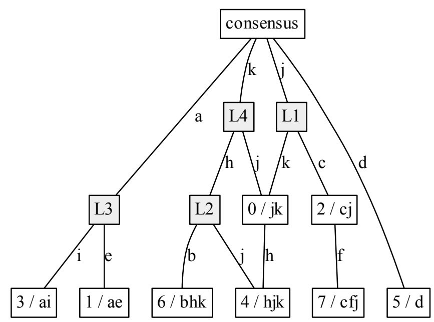
Since data is a binary table, Formal Concept Analysis can be applied. The concept lattice obtained from the data is shown in Fig. 1 (Lower Left). In general, it does not correspond to a median graph. To build a median graph, a necessary condition is to have a distributive $\vee$-semi-lattice. In [9], based on the work of Birkhoff and FCA formalism, we propose an algorithm to compute such a semi-lattice (and the corresponding data table). The result of this algorithm, transforming a concept lattice into a median graph, is given in Fig. 1 (Lower Right). Since FCA is supported by a wide community, the main idea of these researches is to be able to use FCA results and softwares to deal with phylogenetic data and median graphs.

Remark that, for this particular data set the algorithm find the median graph computed by Bandelt in [4], unfortunately, in some cases the algorithm returns a distributive $\vee$-semi-lattice L_d that is not minimal: There exists $L_{d'}$ such that L can be embedded in $L_{d'}$ and $L_{d'}$ can be embedded L_d.

The continuation of [9] is to search for an algorithm which outputs a minimal distributive $\vee$-semi-lattice. Since we look for minimality, a natural question arises: does a unique minimal (so, minimum) distributive $\vee$-semi-lattice L_d exists? In this paper, we propose a counter-example, and then we show that a minimum distributive $\vee$-semi-lattice does not always exist.

In the following section, we recall definitions and notation for the understanding of this paper. We then sketch the limitations of our algorithm and show in Section 4 that a minimum distributive $\vee$-semi-lattice does not exist. We conclude this paper by some remarks and perspectives in Section 5.

	a	b	c	d	e	f	g	h	i	j	k
0							×			×	×
1	×				×						
2			×							×	
3	×								×		
4							×	×		×	×
5				×							
6		×						×			×
7			×			×				×	



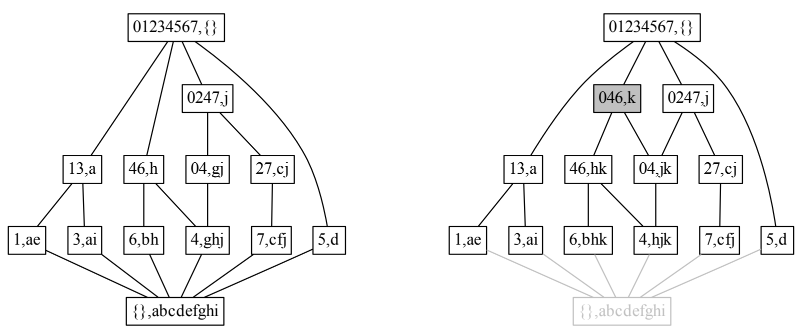


Fig. 1. *Upper Left.* Phylogenetic data of sequence variations for individual groups. *Upper Right.* Median graph obtained from the data ([4]). *Lower Left.* Concept lattice for phylogenetic data. *Lower Right.* Concept lattice corresponding to the median graph. The new concept $(046, k)$ corresponds to $L4$. To obtain this lattice, data must be modified replacing column g by k.

2 Models: lattices, semi-lattices, median algebras and median graphs

In this section we recall basic notions and notation needed throughout the paper. We will mainly adopt the formalism of [8], and we refer the reader to [6,7] for further background. *In this paper, all sets are supposed to be finite.*

2.1 Lattices and FCA

$(\mathcal{J}(L), \mathcal{M}(L), \leq)$	a	b	c
1	×	×	
2		×	
3			×

$(\mathcal{J}(L), \mathcal{J}(L), \ngeq)$	$1(c)$	$2(d)$	$3(b)$
1		×	×
2			×
3	×	×	

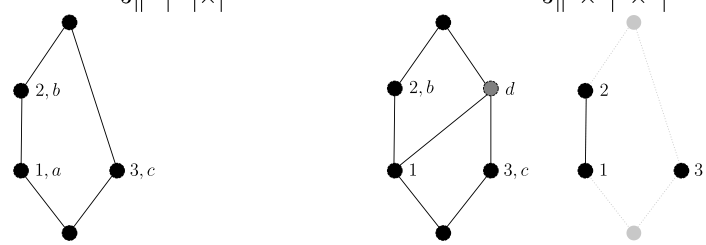


Fig. 2. *Upper Left.* Standard context for lattice N_5 *Lower Left.* N_5, a non distributive lattice. *Upper Right.* The context $(\mathcal{J}(N_5), \mathcal{J}(N_5), \ngeq)$ of an ideal (and so, distributive) lattice. *Lower Right.* Ideal lattice for $(\mathcal{J}(N_5), \leq)$ and this poset. Note that N_5 can be order-embedded in this lattice. Concepts (X,Y) are maximal rectangles of the contexts. For an element e of the lattice, the corresponding concept (X, Y) is $X = \{j \in \mathcal{J}(L) \mid j \leq e\}$ and $Y = \{m \in \mathcal{M}(L) \mid m \geq e\}$. For example, the element with label d is the concept $(13, d)$.

A *partially ordered set* (or *poset* for short) is a pair $(P, \leq)$ where P is a set and $\leq$ is a *partial order* on P, that is, a reflexive, antisymmetric and transitive binary relation on P.

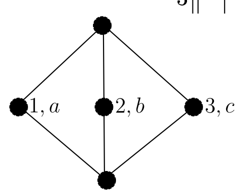
An upper (*resp. lower*) bound of $X \subseteq P$ is an element $y \in P$ such that $\forall x \in X,\ x \leq y$ (*resp* $y \leq x$). For $X \in P$, the lowest upper bound, if exists, is called the join or the supremum. The greatest lower bound, if exists, is called the meet or the infimum.

A $\vee$-semi-lattice $(L, \leq)$ (resp. $\wedge$-semi-lattice) is an ordered set such that the supremum (*resp. infimum*) exists for all $X \subseteq L$. A lattice $(L, \leq)$ is an ordered set such that a supremum and an infimum exist for all $X \subseteq L$. For $x, y \in L$, $x \vee y$ denotes the supremum and $x \wedge y$ denotes the infimum. $\vee$ and $\wedge$ can be considered as binary operators on elements of L. For a finite lattice $(L, \leq)$ there exists $\bot = \bigwedge L$ the lowest element (bottom) and $\top = \bigvee L$ the greatest element (top) of L.

An element $x \in L$ such that $x = y \vee z$ implies $x = y$ or $x = z$ is called a $\vee$-irreducible element. Dually, an element $x \in L$ such that $x = y \wedge z$ implies $x = y$ or $x = z$ is called a $\wedge$-irreducible element. We will denote the set of $\wedge$-irreducible elements and $\vee$-irreducible elements of L by $\mathcal{M}(L)$ and $\mathcal{J}(L)$, respectively. Observe that both $\mathcal{M}(L)$ and $\mathcal{J}(L)$ are posets when ordered by $\leq$.

Posets and lattices can be represented and visualized by their Hasse-diagrams [7]. Examples of lattices are given in Fig. 2 and Fig. 3. Note that some particular lattices have been a name. This is the case for N_5 and M_3, involved in non distributivity (see section 2.2). In Figures 2 and 3, $\vee$-irreducible elements are labeled with numbers and $\wedge$-irreducible elements are labeled with letters. Some elements, doubly irreducibles, have two labels. In Fig. 3 elements 1 and 2 are $\vee$-irreducibles and $\wedge$-irreducible (labeled a and c), elements d and e are $\wedge$-irreducible and element 2 is $\vee$-irreducible.

$(\mathcal{J}(M_3), \mathcal{M}(M_3), \leq)$	a	b	c
1	$\times$		
2		$\times$	
3			$\times$



$(\mathcal{J}(L), \mathcal{M}(L), \leq)$	a	c	d	e
1	$\times$		$\times$	
2			$\times$	$\times$
3		$\times$		$\times$

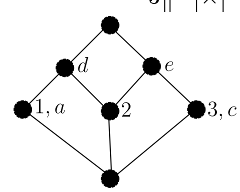


Fig. 3. *Upper Left.* the standard context for lattice M_3. *Lower Left.* M_3 is a non distributive lattice. *Upper Right.* the standard context $(\mathcal{J}(L), \mathcal{M}(L), \leq)$ for L below. *Lower Right.* A non distributive lattice L such that $(\mathcal{J}(L), \leq) = (\mathcal{J}(M_3), \leq)$.

Formal Concept Analysis [8] uses *concept lattices* for data analysis tasks. Concept lattices are built from a binary table, which is called a *formal context*, by the way of the *Galois connection*.

We denote by (G, M, I) a formal context where G is a set of objects, M a set of attributes and I an incidence relation between objects and attributes. In phylogenetic data, objects are usually species, attributes are mutations, and $(g, m) \in I$ –or gIm– indicates that mutation m is spotted in specie g.

Definition 1 (Galois connection). *For a set $X \subseteq G$, $Y \subseteq M$ we define:*

$$X' = \{y \in M \mid xIy \textit{ for all } x \in X\}$$
$$Y' = \{x \in G \mid xIy \textit{ for all } y \in Y\}$$

Then a formal concept is a pair (X, Y) where $X \subseteq G$, $Y \subseteq M$ and $X' = Y$ and $Y' = X$. X is the extent and Y is the intent of the concept. They are closed sets as they verify $X = X''$ and $Y = Y''$. The set of all formal concepts ordered by inclusion of the extents –dually the intents– denoted by $\leq$ generates the *concept lattice* of the context (G, M, I). The existence of a supremum and an infimum allows to use lattices for classification process. Concepts can be viewed as classes, indeed a concept (X, Y) is a representation of a maximal set of objects X which share a maximal set of attributes Y. If another concept (X_1, Y_1) is greater than (X, Y), it contains more objects, but described by fewer attributes. X_1 can be considerated as a class, more general than X.

A *clarified context* is a context such that $x' = y'$ implies $x = y$ for any element of G and any element of M. In a clarified context, the set of attributes of two distinct objects are distincts, and dually for objects. Moreover, a clarified context is *reduced* iff it contains:

– no vertex $x \in G$ such that $x' = X'$ with $X \subseteq G$, $x \notin X$
– no vertex $x \in M$ such that $x' = X'$ with $X \subseteq M$, $x \notin X$

Indeed, a vertex $x \in G$ such that $x' = X'$ with $X \subseteq G$, $x \notin X$ correspond to a irreducible element (since it may be reduced to others elements by galois connection). Only irreducible elements, which are not join or meet of others elements, are necessary to build a lattice. The reduced context is also called a *standard context* [8]. Note that the standard context of lattice L is such that $G = \mathcal{J}(L)$ and $M = \mathcal{M}(L)$. Examples of standard contexts are given in Fig. 2 and Fig. 3. The corresponding concept lattice is given below the context.

2.2 Distributive lattices

As stated in the motivations, median graphs are used for phylogenetic purposes, and encode a family of trees. It is known that theses graphs can be considered as particular distributive $\vee$-semi-lattices. This subsection provides basic notions about distributive (semi)-lattices.

A lattice is *distributive* if $\wedge$ and $\vee$ are distributive one with respect to the other. Formally, a lattice L is distributive if for every $x, y, z \in L$, we have that one (or, equivalently, both) of the following identities holds:

$$(i)\ x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z), \quad (ii)\ x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

Distributive lattices appear naturally in any classification task or as computation and semantic models; see, e.g., [6,7,10,11]. This is partially due to the fact that *any distributive lattice can be thought of as a sublattice of a power-set lattice, i.e., the set $\mathcal{P}(X)$ of subsets of a given set X.*

Note that the definition of a sublattice is more constraint than the definition of a suborder: A subset $X \subseteq L$ is a *sublattice* of L if for every $x, y \in X$ we have that $x \wedge y, x \vee y \in X$. For example, in Fig. 2, N_5 (left) is a suborder of the lattice on the right, but is not a sublattice. Indeed, $1 \vee 3$ is not the same element in the two lattices.

The distributivity property of lattices has been equivalently described in several ways. One of these properties relies the notion of sublattice, as follows:

Property 1. L is a distributive lattice iff it does not contain neither N_5 nor M_3 as sublattices.

This property describes distributive lattices in terms of two forbidden structures, namely, M_3 and N_5 that are, up to isomorphism, the smallest non distributive lattices. N_5 is represented in Fig. 2 (Lower left) and M_3 is represented in Fig. 3 (Lower left). In Fig. 2 the lattice at the right corner does not contain N_5 nor M_3 as sublattices, and so is distributive. In Fig. 3 the lattice at the right corner does not contain M_3 as sublattice, but contains N_5 as a sublattice, and so, is not distributive.

Those properties of distributive lattice are useful to check whether a lattice is distributive or not. Our goal is to transform a lattice into a distributive one. For this particular task, the Birkhoff representation of distributive lattice is of practical interest. It use the notion of *order ideal*, recalled here:

Definition 2 (Order Ideal). *Let* $(P, \leq)$ *be a poset. For a subset* $X \subseteq P$, *let* $\downarrow X = \{y \in P : y \leq x \text{ for some } x \in X\}$ *and* $\uparrow X = \{y \in P : x \leq y \text{ for some } x \in X\}$. *A set* $X \subseteq P$ *is a (*poset*)* ideal *(resp.* filter*) if* $X = \downarrow X$ *(resp.* $X = \uparrow X$*). If* $X = \downarrow \{x\}$ *(resp.* $X = \uparrow \{x\}$*) for some* $x \in P$, *then* X *is said to be a* principal *ideal (resp. filter) of* P. *For principal ideals, we omit brackets, so that* $\uparrow x$ *(resp.* $\downarrow x$*) stands for* $\uparrow \{x\}$ *(resp.* $\downarrow \{x\}$*)*

$(\mathcal{J}(L), \mathcal{J}(L), \not\geq)$	$1(f)$	$2(e)$	$3(d)$
1		×	×
2	×		×
3	×	×	

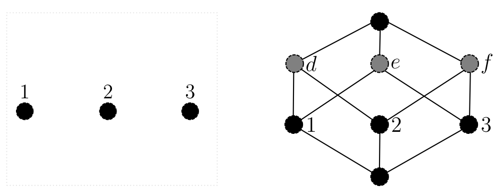


Fig. 4. *Left.* The context $(\mathcal{J}(L), \mathcal{J}(L), \not\geq)$ for $(J(L), \leq)$ the poset induced by $\vee$-irreducible elements of lattices in Fig. 3. *Middle.* $(J(L), \leq)$ for L (or equivalently M_3 in Fig. 3. *Right.* Ideal lattice for $(\mathcal{J}(L), \leq)$ (equivalently $(\mathcal{J}(M_3), \leq)$). M_3 and L can be order-embedded in this lattice.

Birkhoff's representation of distributive lattices.

Let $(P, \leq)$ be a poset and consider the set $\mathcal{O}(P)$ of ideals of P, i.e.,

$$\mathcal{O}(P) = \{\bigcup_{x \in X} \downarrow x \mid X \subseteq P\}.$$

It is well-known that for every poset P, the set $\mathcal{O}(P)$ ordered by inclusion is a distributive lattice, called *ideal lattice* of P. Furthermore, the poset of $\vee$-irreducible elements of $\mathcal{O}(P)$ is $\mathcal{J}((\mathcal{O}(P))) = \{\downarrow x \mid x \in P\}$ and it is (order) isomorphic to P.

This representation is used to provide a distributive lattice L_d with the same poset of $\vee$-irreducible elements as an arbitrary lattice L. In this case, L is order-embedded in L_d. For example, in Fig. 2, the lattice on the right corner is the ideal lattice of $(\mathcal{J}(N_5), \leq)$. In the same way, the two lattices on the left in Fig. 3 have the same poset of $\vee$-irreducible elements, and can be embedded in the ideal lattice of this poset, represented in Fig. 4. In particular, in [12,5] it is shown that the family of lattices with the same poset of $\vee$-irreducible elements is itself a lattice, an so there exists a minimum element.

From a poset P, it is possible to obtain the context of the ideal lattice as $C = (P, P, \not\geq)$. For a standard context $C = (\mathcal{J}(L), \mathcal{M}(L), \leq)$, the standard context of the ideal lattice is $C = (\mathcal{J}(L), \mathcal{J}(L), \not\geq)$. Note also that, for every distributive lattice L, the two posets $\mathcal{J}(L)$ and $\mathcal{M}(L)$ are dually-isomorphic. This is why the standard contexts of distributive lattices are "squares" ($|\mathcal{J}(L)| = |\mathcal{M}(L)|$), and are built with the information of only one of these two posets.

In the following subsection, we give some hints about median graphs and median algebras. As we will soon observe, the class of median graphs is in correspondance with a particular subclass of distributive $\vee$-semi-lattices.

Let L be a $\vee$-semi-lattice and $x \in L$, then $\uparrow x$ is a lattice (in the finite case, every $\vee$-semi-lattice with a lowest element is a lattice). Then a $\vee$-semi-lattice L is distributive iff $\uparrow x$ is distributive, for all x [6]. In practice, it is sufficient to check this property only for minimal elements of L. Indeed, filters of non minimal elements are sublattices of a minimal element filter, and sublattices of distributive lattices are distributives.

2.3 Median graphs

As said in the introduction, a median graph encodes all parcimonious phylogenetic trees. A median graph is a connected graph having the median property, *i.e.* for any three vertices a, b, c, there is exactly one vertex x which lies on a shortest path between each pair of vertices in $\{a, b, c\}$.

The following characterization of distributive lattices explains some links of distributive lattice with median graphs and median algebras.

Property 2. A lattice L is a distributive lattice iff for all $x, y, z \in L$,

$$(x \wedge y) \vee (y \wedge z) \vee (z \wedge x) = (x \vee y) \wedge (y \vee z) \wedge (z \vee x).$$

This property establishes a correspondence between distributive lattices and median algebras. Indeed, a median algebra is a structure (M, m) where M is a nonempty set and $m : M^3 \rightarrow M$ is an operation, called *median operation*, that satisfies the following conditions $m(a, a, b) = a$ and $m(m(a, b, c), d, e) = m(a, m(b, c, d), m(b, c, e))$, for every $a, b, c, d, e \in M$. It is not difficult to see that if L is distributive, then $m(a, b, c) = (a \wedge b) \vee (b \wedge c) \vee (c \wedge a)$ is a median operation. The connection to *median graphs* was established by Avann [1] who showed that every median graph is the Hasse diagram of a median algebra (thought of as a semilattice). For further background on median structures see, e.g., [2]. This result was later used by Bandelt [3] to establish the link between distributive lattices and median graphs.

Property 3. A graph is a median graph iff it is isomorph to a $\vee$-semi-lattice L with the two following properties:

- L is distributive
- for all $x, y, z \in L$ such that $(x \wedge y)$, $(y \wedge z)$ and $(z \wedge x)$ are defined, $(x \wedge y \wedge z)$ is defined.

3 Algorithm to produce a distributive $\vee$-semi-lattice

To build a median graph from a context using FCA, a necessary condition is to build a distributive $\vee$-semi-lattice. For the concept lattice L, it is always possible to consider the semi-lattice $L_\vee = L \backslash \bot$ (L minus the lowest element). Minimal elements of this semi-lattice $L_\vee$ are minimal elements of $(\mathcal{J}(L), \leq)$.

It remains to transform the filter of these elements into a distributive lattice. Our previous work [9] is based on Birkhoff's representation of a distributive lattice. Since sublattices of a distributive lattice are distributive, a simple way to obtain a distributive $\vee$-semi-lattice from a lattice L is to map L into the ideal lattice of $(J(L), \leq)$. In practice, the bottom element exists because of the existence of infimum in lattice, but it usually does not have semantic for classification. For example, trees are median graphs and so distributive $\vee$-semi-lattice (considering the root as the greatest element) but obviously not lattices. With the adjonction of a bottom element $\bot$, the trees become lattices. There is no reason that these lattices are distributive. Two trivial examples are N_5 and M_3: Once the lowest element is removed, either $N_{5\vee} = N_5 \backslash \bot$ and $M_{3\vee} = M_3 \backslash \bot$, considered as $\vee$-semi-lattices, are distributive (N_5 and M_3 are isomorphic to path and tree). Nevertheless, neither N_5 nor M_3 are distributive.

Now, the mapping of the concept lattice into its ideal lattice will produce a $\vee$-semi-lattice, but this is not necessarily a minimal solution. For example, M_3 will be embedded in the boolean lattice while $M_3 \backslash \bot$ is already a distributive $\vee$-semi-lattice. Hence, the global approach that embeds a concept lattice into its ideal lattice is not efficient.

Alternatively, we can think of a local approach: instead of embedding the whole concept lattice into its ideal lattice, we do so for the sublattices corresponding to filters of minimal elements of $\mathcal{J}(L)$. The algorithm proposed in [9] computes contexts of $\uparrow j$ for every minimal $\vee$-irreducible element j, and transforms these contexts so that they correspond to the context of a distributive lattice. Once these contexts are built, we merge them to build the whole lattice.

However, this method does not always output a minimal solution, *i.e.*, there may exist $L_{d'}$ a distributive $\vee$-semi-lattice such that L can be embedded in $L_{d'}$, and $L_{d'}$ can be embedded in L_d with $|L| < |L_{d'}| < |L_d|$ and $(\mathcal{J}(L), \leq) = (\mathcal{J}(L_{d'}), \leq) = (\mathcal{J}(L_d), \leq)$. This result comes from the fact that each filter is processed independently. Nevertheless, it is possible that some elements are

Algorithm 1: Construction of context of a distributive $\vee$-semi-lattice.

Data: A context $(\mathcal{J}(L), \mathcal{M}(L), I)$ of a lattice L

Result: the context $(\mathcal{J}(L_{med}), \mathcal{M}(L_{med}), I)$ of a distributive $\vee$-semi-lattice L_{med} such that L can be order-embedded in L_{med}

foreach $j \in \mathcal{J}(L)$, *minimal* **do**
  $(P_j, \leq) \leftarrow \emptyset$

repeat
  stability $\leftarrow$ true;
  foreach $j \in \mathcal{J}(L)$, *minimal* **do**
    compute P_j the poset of $\vee$-irreducible elements in $\uparrow j$
    compute $C_j = (P_j, P_j, \not\geq)$
    if P_j *modified since last iteration* **then**
      stability $\leftarrow$ false;
  Merge all $C_j = (P_j, P_j, \not\geq)$ in a unique context
  Reduce this context
until *stability*

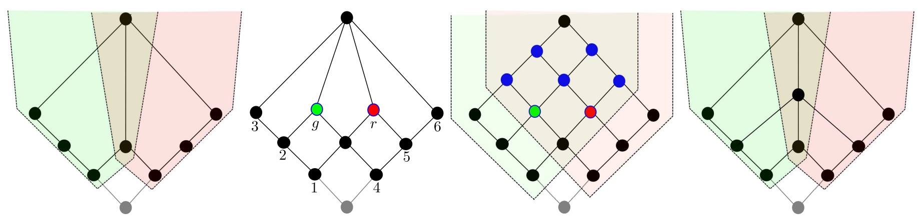


Fig. 5. *From left to right:* A lattice. Result of the first step of the algorithm. Result of the algorithm. A minimal distributive $\vee$-semi-lattice (not reachable by the algorithm).

shared by several filters of minimal $\vee$-irreducible elements. This is illustrated in Fig. 5, and motivates the two following observations.

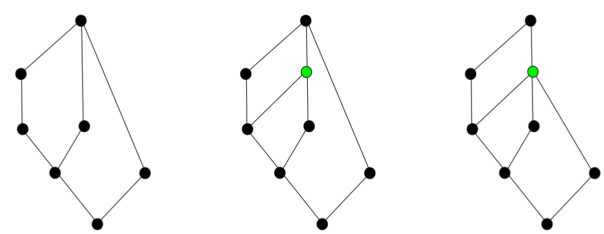
First, it is possible that some elements added to a filter for achieving distributivity belong to others filters. These new elements may break a previously obtained distributivity in others filters. This is the case in Fig. 5 with the two new elements (the red one and the green one). At a first iteration of the loop of the algorithm, when the filters are merged, g and r are distinct elements. Neither the filter of 1 nor 2 are distributive. $\uparrow 1$ (resp. $\uparrow 2$) is not distributive because of r (resp. g). To overcome the problem, the algorithm loops while any filter is modified by the process. At worst, the algorithm computes the context corresponding to the ideal lattice of $\vee$-irreducible poset of L and the algorithm always terminates.

Second, in some cases, a minimal solution cannot be reached when locally considering the filters. Such a solution is proposed if Fig. 5 (extreme right)

4 A counter-example for the existence of a minimum distributive $\vee$-semi-lattice

The local approach thus seems to be better than the global one. However, our algorithm does not always produce a minimal solution. The natural question is then whether, for a lattice L, there exists a minimum (*i.e.*, minimal and unique) distributive $\vee$-semi-lattice L_d such that L can be embedded into L_d. We will now show through a counter-example that such minimum does not always exist.

The proposed counter-example is given in Fig. 6: For the lattice shown in (a), either lattice in (b) and (c) are minimal distributive $\vee$-semi-lattices (since they differ by one element only) but it is obvious that (b) and (c) are not isomorphic. So, since a minimum solution does not exist, some choices remain to do in goal to use FCA algorithms for traditional application fields of median graphs, in particular for phylogeny.



(a) (b) (c)

Fig. 6. A lattice (a) such that there exists two non isomorphic minimal distributive $\vee$-semi-lattices (when removing bottom element) (b) and (c)

5 Discussion and perspectives

We have seen that there is a lattice L for which there is not a unique minimum distributive $\vee$-semi-lattice L_d such that L can be embedded in L_d and with the same posets of $\vee$-irreducible elements $((\mathcal{J}(L), \leq) = (\mathcal{J}(L_d), \leq))$.

So, even if we provide an algorithm that produces a minimal solution, the question of the meaning of this (not unique) solution should be addressed. A way to tackle it is to find an algorithm able to list all the minimal solutions. Alternatively, we could propose a measure of "interestingness" of these minimal solutions, so that an optimal solution could be reached based on such a measure. This remains a topic of current research.

Also, this work was motivated by the study of the relations between distributive $\vee$-semi-lattices and median graphs. Not all distributive $\vee$-semi-lattices are median graphs. It remains to check the following condition: For every triple of elements x, y, z such that $x \wedge y$, $x \wedge z$ and $y \wedge z$ are defined, $x \wedge y \wedge z$ is defined.

It is obvious that this condition is not satisfied for some distributive $\vee$-semi-lattices. A trivial example is the Boolean lattice (minus the bottom element) but in this particular case, the whole lattice is distributive, and so a median graph. Nevertheless, it remains open whether this is always the case.

References

1. Avann, S.P.: Median algebras. Proceedings of the American Mathematical Society **12**, 407–414 (1961)
2. Bandelt, H.J., Hedlíková, J.: Median algebras. Discrete mathematics **45**(1), 1–30 (1983)
3. Bandelt, H.J., Forster, P., Röhl, A.: Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution **16**(1), 37–48 (1999)
4. Bandelt, H.J., Macaulay, V., Richards, M.: Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtdna. Molecular phylogenetics and evolution **16**(1), 8–28 (2000)
5. Bordalo, G.H., Monjardet, B.: The lattice of strict completions of a poset. Electronic Notes in Discrete Mathematics **5**, 38–41 (2000)
6. Caspard, N., Leclerc, B., Monjardet, B.: Finite ordered sets: concepts, results and uses. Cambridge University Press (2012)
7. Davey, B.A., Priestley, H.A.: Introduction to Lattices and Order. Cambridge university press (2002)
8. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer (1999)
9. Gély, A., Couceiro, M., Napoli, A.: Steps towards achieving distributivity in formal concept analysis. In: Proceedings of the Fourteenth International Conference on Concept Lattices and Their Applications, CLA 2018, Olomouc, Czech Republic, June 12-14, 2018. pp. 105–116 (2018), http://ceur-ws.org/Vol-2123/paper9.pdf
10. Hopcroft, J.E., Motwani, R., Rotwani, Ullman, J.D.: Introduction to Automata Theory, Languages and Computability. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edn. (2000)
11. Mattern, F.: Virtual time and global states of distributed systems. Parallel and Distributed Algorithms **1**(23), 215–226 (1989)
12. Nation, J., Pogel, A.: The lattice of completions of an ordered set. Order **14**(1), 1–7 (1997)
13. Priss, U.: Concept lattices and median networks. In: CLA. pp. 351–354 (2012)
14. Priss, U.: Representing median networks with concept lattices. In: ICCS. pp. 311–321. Springer (2013)
15. Vigilant, L., Pennington, R., Harpending, H., Kocher, T.D., Wilson, A.C.: Mitochondrial dna sequences in single hairs from a southern african population. Proceedings of the National Academy of Sciences **86**(23), 9350–9354 (1989)